Integrating Statistics and Geography to climate change vulnerability studies

Maria do Carmo Dias Bueno *
IBGE, Rio de Janeiro, Brazil – maria.bueno@ibge.gov.br

Álvaro de Oliveira D'Antona
UNICAMP, Campinas, Brazil – alvaro.dantona@fca.unicamp.br

Abstract

We are all aware that the impacts of climate change on human populations are more and more frequent and potentially devastating, therefore we need more effective actions to understand our vulnerability to extreme climatic events. The deep comprehension of the phenomena is the key to deal with them and its consequences.
There are clearly two dimensions within this discussion which need to be integrated from a spatial perspective: the social and the environmental dimensions. In a broad way, the social dimension could be represented by statistics, particularly census data, and the environmental dimension could be represented by susceptibility maps. The exploitation of demographic data could provide a wealthy analysis in which can guide more effective approaches, particularly, when it is applied to the geography of climate related hazards.
First of all, there are some methodological challenges to be overcome.
Historically, statistics has been released to administrative boundaries (country, states, provinces, municipalities) and, in a few cases, to some operational boundaries (enumeration areas, blocks, output areas). The climate hazards, however, operate cutting across these boundaries. So, statistics need to be processed in reference to the geography of climate hazards. In most of the cases, the statistical granularity level data is quite low and it is also necessary a size reduction of the aggregation units. Nowadays, these tasks are not hard to perform in light of the technological advances in GIS and its growing use in National Statistical Offices.
We show a study case in Brazil which uses census microdata aggregate in small and regular geographical units that are used to characterize the population settled in susceptible areas to geological hazards and flooding at the northern coast of São Paulo.
We could conclude that census data integrated to geography of climate change and connected processes can significantly improve the understanding of hazards and vulnerability. Until this result can be reached, there are many steps ahead. The presence of geo-referenced data is not enough, neither statistical nor geographical data. Thus, it is necessary a common spatial basis for both, so that the data integration can be simple, quick, and efficient.

Keywords: geography; census; data integration; gridded data.

1. Introduction

In a short period of time, the impacts of climate change have become the greatest challenge in the twenty-first century, and its consequences are felt by the most part of the population. There are no locations free of risks and there is always space in the newspaper headlines to announce extreme weather events.
The effects on human health are increasingly present; changes in precipitation frequency and intensity patterns affect agriculture and lead to flooding in some places and drought in others; landslides, sea level rise and storms are more frequent. What yesterday was wondered to be science fiction, nowadays is the almost certain future of humanity (IPCC 2014).
If we consider that most of the population lives in urban areas, we can say that big cities will have the greatest impact of climate changes (Kowarick 2002; Kasperson et al. 1995). These impacts will be

further intensified by environmental problems typical of large urban areas where the environmental infrastructure is inconsistent with the population that consumes its resources.
In the context of a research agenda on population and climate change, is essential there be an effort to develop methodologies and indicators for the identification and characterization of the areas at risk due to climate change and the vulnerable populations, especially in urban areas.

2. Regular Grids

Since the late 1980s, there has been a growing interest in the "human dimensions of environmental change" looking for means to integrate methodologies and data coming from the natural sciences and from the social sciences (Liverman et al. 1998).
But these subjects often use different units of analysis, since the phenomena and characteristics of the biophysical environment are measured in a not compatible form with the phenomena and socio-demographic characteristics. Data collected by population and housing censuses, for example, are historically oriented to administrative units. However, climate events occur in different scales and units, which never are associated with such administrative units.
The starting point for understanding the human dimensions of environmental change should be a common spatial support to allow data integration, once most environmental events do not occur under the same spatial and temporal basis that the socioeconomic events occur. In this way, finding a common basis that makes the data integration easier and that allow researchers to work independently of political or administrative units is an important step (Hogan & Marandola Jr. 2012).
In recent decades, technological advances have allowed a knowledge improvement of the relationship between environment and population, mainly by the incorporation of geotechnologies and by the use of space as an integration plan (Cortez & D'Antona 2014).
At first, this challenge was partly supplied with dasimetric techniques to generate estimates of total population using ancillary data, for instance, the Center for International Earth Science Information Network – CIESIN developed the Gridded Population of the World – GPW (1995) and the Global Rural Urban Mapping Project – GRUMP (2004). Currently, other variables are estimated using modern techniques and detailed data, what on one hand, has advantages such as production of annual estimates, but on the other hand, these data are highly dependent on the ancillary data quality (Tatem et al. 2007; Tatem et al. 2014).
Although census data are disseminated using large units in some countries, making harder realizing detailed analysis, they have a high geographical coverage. Nevertheless, its use on studies of adaptation to climate change has been too small and its potential has been underestimated (Guzmán et al. 2013).
Since the censuses round 2000, there has been an increasing use of geotechnologies in all stages of the census operations, which made possible to obtain the results for various geographies besides the traditional ones. One approach that has been used by several countries, initially from the Northern Europe (EFGS 2012), and more recently by others, such as Spain and Brazil (INE-Spain 2014; Bueno 2013), is the use of a 1 km^2 cell (or smaller[1]) arranged in a structure of regular grids for aggregation of census data. This approach allows grid cells to be joined to shape any desired geographical clipping and it also allows them to be aggregated into larger cells to meet the needs of assessments in different scales (Bueno 2014).
The methodology for generating this data can be aggregation, disaggregation or hybrid. The first method uses the census microdata associated with their locational attributes for data aggregation; the second uses spatial and/or statistical methods with ancillary data for the spatial data shift; the third combines the two previous ones, according to the existence of spatial information (Bueno 2014).
Although the census happens every 10 years only, its benefit is a regular and harmonized production of a minimum set of variables to virtually every country. The variables obtained in the census provide

[1] The issue of the confidentiality is the major challenge in putting census data aggregated in small geographical units, since the decrease of the area of the aggregation unit increases significantly to the disclosure risk.

minimally information about the quantity, location and main characteristics of households; with respect to population, the censuses provide information on gender and age. These are the basic information necessary for determining the risks associated with environmental issues. To meet the need for more detailed variables there is the possibility of combining the census with local sample surveys or administrative data (Guzmán et al. 2013).

3. Case study

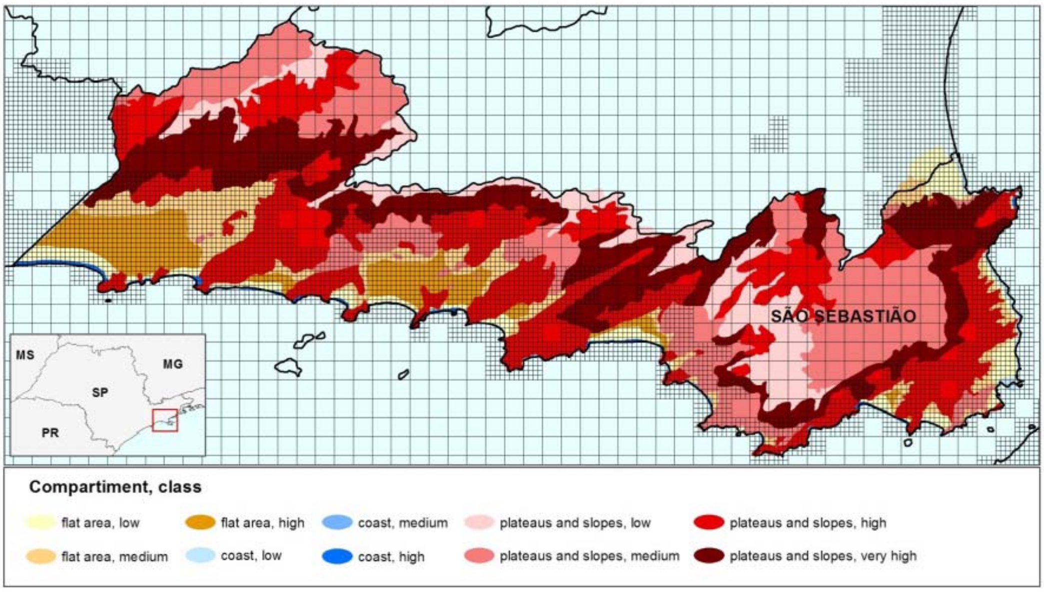
The following case study intends to illustrate the data integration to assist studies and analysis of the interaction between the human dimension and the environmental processes.
São Sebastião was the site selected for the integration of susceptibility data to mass movements and flood with demographic data related to some characteristics of the population and households. This municipality is located on the northern coast of São Paulo and is inserted into a region of great economic importance due to the concentration of ports, oil, and gas pipelines. Moreover, it is a tourist center that attracts lots of people on weekends and summer holidays.
Often there are natural disasters involving landslides in this municipality, mainly because their geophysical characteristics and disordered population growth, which led to the settlement of areas near to the slopes and hills.
The applied techniques in this case are suited to other contexts and geographical clippings and can be easily reproduced.
As indicated above, the grid with demographic data was generated by combining techniques of aggregation and disaggregation. The cell dimensions are 1 x 1 km in rural areas and 200 x 200 m in urban areas (Bueno 2014). The demographic variables used in this data integration were: total population, gender, age, household income, type of property occupancy, and type of household.
The classification of susceptibility has considered the specificities of each physiographic compartment, i.e., in the plateaus and slopes were analyzed the mass movements processes; in the flat areas were analyzed the flooding susceptibility; and in the coast was considered the coastal erosion (Geological Institute 1996).
The processes susceptibility class polygons were overlaid by the grid cells containing the census data (Figure 1) to allow the execution of a spatial operation, which result give us the volume of population and their demographic characteristics.
The results have showed us that less than 2% of the population lives in areas of susceptibility what is considered very high, and that about 40% of the population lives in high susceptible areas. Children up to 9 years old represent a significant volume in this portion of the population - about 6%. Concerning to household income, the highest amount of households who have reported having any income and those in the lower income range are located in high susceptibility areas. These data confirm that the less privileged socioeconomic groups are more vulnerable, since their homes are usually located in poor urbanization areas (Hogan & Marandola Jr. 2012).
Regarding to the other variables, there were no significant frequencies related to the susceptibility classes.

Table 1 – Percentage of population according to age group by susceptibility class.

susceptibility	0 to 9 y.o. (%)	10 to 19 y.o. (%)	20 to 64 y.o. (%)	more than 65 y.o. (%)	Total (%)
very high	0.21	0.27	0.82	0.09	1.39
high	6.24	6.70	21.87	1.26	36.07
medium	4.77	5.37	19.25	1.80	31.19
low	4.27	5.12	19.61	2.34	31.36

Source: Instituto Geológico 1996; IBGE 2010. Prepared by the authors.

Figure 1 – Susceptibility map overlaid by regular grids.



Source: Instituto Geológico 1996. Prepared by the authors.

Table 2 – Percentage of population according to household income in minimum wage[2] ranges by susceptibility class.

susceptibility	without income (%)	up to 3 (%)	3 to 5 (%)	5 to 10 (%)	more than 10 (%)
very high	0.07	1.16	0.08	0.03	0.01
high	1.15	31.33	1.45	0.95	0.24
medium	0.93	26.63	1.91	1.24	0.37
low	0.75	26.63	2.84	1.75	0.49

Source: Instituto Geológico 1996; IBGE 2010. Prepared by the authors.

4. Conclusions

The results for the municipality of São Sebastião have not showed us any surprises, and it confirms what is a universal knowledge: the most vulnerable to climate change is also part of a social and economically vulnerable group.

Knowing that environmental changes caused by climate change can no longer be reversed, it is necessary to focus on the understanding and reduction of the vulnerabilities so that we can reach an effective adaptation to this new reality. The knowledge of the population exposed to risks and hazards associated with climate change is part of this process of understanding as well as the understanding of the mechanisms that lead to the exposition to this risks and hazards.

In order to do that, it is necessary to develop methods to make easy the joining of data from different scales and from different fields of the knowledge aiming to turn multi-scale and multi-interdisciplinary analysis a feasible task.

But only the existence of geo-referenced data is not enough, it is necessary to have a layer that serves as a spatial support for the all data aggregation.

The system of regular grids seems to be an ideal spatial support to do that.

[2] The minimum wage in Brazil in 2010 was R$ 510 and was equivalent to US$ 290 (exchange rate at that time).

References

Bueno, M.C.D. (2013). Utilização de grades em estatística. Proceedings of the 2nd Methodology Seminar of IBGE. http://eventos.ibge.gov.br/images/smi2013/downloads/ST6_Utilizacao_de_grades_em_estatistica.pdf

Bueno, M.C.D. (2014). Grade estatística: uma abordagem para ampliar o potencial analítico de dados censitários. PhD Thesis, Instituto de Filosofia e Ciências Humanas, UNICAMP.

Côrtes, J.C. & D'Antona, A.O. (2014). Dinâmicas no uso e cobertura da terra – perspectivas e desafios da Demografia. *Revista Brasileira de Estudos de População*, **31**.

Guzmán, J.M., Schensul, D. & Zhang, S. (2013). Understanding Vulnerability and Adaptation Using Census Data. In: Martine, G. & Schensul, D. (Eds.). *The Demography of Adaptation to Climate Change*. New York, London and Mexico City: UNFPA, IIED and El Colegio de México.

Hogan, D. & Marandola Jr., E. (2012). Bringing a population-environment perspective to hazards research. *Population & Environment*, **34**.

Instituto Geológico. (1996). Carta de risco a movimentos de massa e inundação do município de São Sebastião.

INE Spain. (2014). *Indicators for grids of 1 km^2*. http://www.ine.es/en/censos2011_datos/cen11_datos_resultados_rejillas_en.htm

IPCC. (2014). *Climate Change 2014: Synthesis Report. Summary for Policymakers.* Geneva, Switzerland: IPCC.

Kasperson, J.X., Kasperson, R.E. & Turner II, B.L. (1995). *Regions at risk: comparisons of threatened environments.* Tokyo: United Nations University.

Kowarick, L. (2002). Viver em risco: sobre a vulnerabilidade no Brasil urbano. *Novos Estudos*, **63**.

Liverman, D. Moran, E.F., Rindfuss, R.R. & Stern, P.C. (1998). *People and pixels: linking remote sensing and social science.* Washington, D.C.: National Academy Press.

Tatem, A.J., Noor A.M., von Hagen C., Di Gregorio A. & Hay, S.I. (2007). High Resolution Population Maps for Low Income Nations: Combining Land Cover and Census in East Africa. *PLoS ONE*, **2**.

Tatem, A.J., Campbell, J., Guerra-Arias, M, de Bernis L., Moran, A. & Matthews, Z. (2014). Mapping for maternal and newborn health: the distributions of women of childbearing age, pregnancies and births. *International Journal of Health Geographics*, **13**.